

# Consequences of local inter-strand dehybridization for large-amplitude bending fluctuations of double-stranded DNA

David A. Sivak<sup>1, a)</sup> and Phillip L. Geissler<sup>1, b)</sup>

*Biophysics Graduate Group, University of California, Berkeley*

(Dated: 1 November 2011)

The wormlike chain (WLC) model of DNA bending accurately reproduces single-molecule force-extension profiles of long (kilobase) chains. These bending statistics over large scales do not, however, establish a unique microscopic model for elasticity at the 1-10 bp scale, which holds particular interest in biological contexts. Here we examine a class of microscopic models which allow for disruption of base pairing (i.e., a ‘melt’ or ‘kink’, generically an ‘excitation’) and consequently enhanced local flexibility. We first analyze the effect on the excitation free energy of integrating out the spatial degrees of freedom in a wormlike chain. Based on this analysis, we present a formulation of these models that ensures consistency with the well-established thermodynamics of melting in long chains. Using a new method to calculate cyclization statistics of short chains from enhanced-sampling Monte Carlo simulations, we compute  $J$ -factors of a meltable wormlike chain (MWLC) over a broad range of chain lengths, including very short molecules (30 bp) that have not yet been explored experimentally. For chains longer than about 120 bp, including most molecules studied to date in the laboratory, we find that melting excitations have little impact on cyclization kinetics. Strong signatures of melting, which might be resolved within typical experimental scatter, emerge only for shorter chains.

PACS numbers: 87.10.Pq, 87.10.Rt, 87.14.gk, 87.15.ak, 87.15.La

Keywords: DNA, wormlike chain, melting, cyclization,  $J$ -factor

## I. INTRODUCTION

The challenge of constructing coarse-grained models, e.g., for large biomolecular systems, lies in accounting for the influence of fluctuations that are not explicitly represented. Systematic procedures for coarse-graining perform such an accounting, often approximately, so that physical consequences of fluctuating solvent densities, electric fields, etc. can be incorporated at low computational cost. In the simplest cases, integrating out certain degrees of freedom just renormalizes interaction parameters for the remaining variables. In general new types of interactions are introduced as well.

This familiar process of renormalization has important implications for the way experimental data should be used to parameterize microscopic models. Measured statistics of a particular variable include the effects of all other fluctuating degrees of freedom. Thermodynamic parameters inferred for that variable do not directly inform the energetics of models that explicitly include other degrees of freedom. In some situations these considerations are transparent: an experimentally derived implicit solvent potential should not be used in models that explicitly represent solvent fluctuations. In other cases the issue can be somewhat more subtle. This paper demonstrates the importance of disentangling renormalization

effects in the specific context of models for bending of nucleic acid molecules.

The wormlike chain (WLC) model has proven remarkably successful in reproducing experiments probing the conformational flexibility of DNA. In a discretized form the model envisions DNA as a fluctuating chain of discrete links, each inextensible in its length<sup>1,2</sup>, with the persistence length  $\ell_P \sim 50$  nm ( $\sim 150$  basepairs [bp]) setting the contour length scale over which orientational correlations decay. The model shows very good agreement with single-molecule force-extension measurements on kilobase-long  $\lambda$ -phage genomic DNA.<sup>1,3</sup> The WLC with  $\ell_P = 50$  nm also accurately predicts cyclization rates of medium-length (hundreds of bp) DNA chains in ligation experiments.<sup>4</sup> In general the WLC model reproduces the results of experiments whose observations are dominated by conformations typical of thermal equilibrium.

Yet in biological contexts DNA is often bent on much shorter length scales than the micron lengths of  $\lambda$ -phage DNA: prokaryotic transcription initiation,<sup>5</sup> nucleosomal genome compaction in eukaryotes,<sup>6</sup> DNA-binding by architectural proteins such as IHF,<sup>7</sup> and viral DNA packaging<sup>8</sup> all feature DNA bending hundreds of degrees on length scales of tens of bp. Furthermore, DNA is increasingly being used as a programmable template for constructing nanomaterials,<sup>9–12</sup> where its bending flexibility over tens of bp can strongly influence the resulting structures.<sup>13</sup> Knowledge of the mechanical properties of short DNA chains is thus essential to understand the role of DNA looping in gene regulation, the nature of protein-DNA interactions, the pressure generated in viral DNA packaging, and the patterned nanomaterial dictated by a given DNA scaffold.

Many microscopic bending potentials produce long

<sup>a)</sup>Current address: Physical Biosciences Division, Lawrence Berkeley National Laboratory

<sup>b)</sup>geissler@berkeley.edu; Department of Chemistry, University of California, Berkeley; Chemical Sciences Division, Lawrence Berkeley National Laboratory

length scale bending statistics identical to the WLC,<sup>14</sup> and thus the relatively well-established long length scale results do not distinguish between substantially different possibilities for smaller length scale bending potentials. In particular, the energetics associated with large end-to-end forces, or equivalently, during exceedingly rare large thermal fluctuations, could take on numerous different forms without substantially changing the long length scale bending statistics.

Indeed, recent experiments have suggested that oligomers shorter than a single persistence length may have considerably different mechanics than a WLC with  $\ell_P = 50$  nm. Measuring the end-to-end distance of DNA free in solution by Förster Resonance Energy Transfer (FRET) and its radius of gyration by small angle X-ray scattering (SAXS), Yuan, *et al.* have inferred an apparent persistence length of 20 nm for chains comprising fewer than 21 bp,<sup>15</sup> though it is not clear how this conclusion is consistent with the accepted value inferred from experiments on longer DNA chains. Cyclization rates of short chains are particularly dominated by rare large fluctuations, and thus should provide a sensitive test of DNA's detailed short length scale bending statistics. Indeed, DNA ligation experiments by Cloutier and Widom found that DNA of approximately 100 bp cyclized as much as five orders of magnitude more readily than expected from corresponding WLCs with  $\ell_P = 50$  nm.<sup>16</sup> By contrast, in similar experiments Du, *et al.* found the cyclization efficiency of small DNA molecules to be in good agreement with the traditional WLC model of DNA bending albeit with a slightly shorter persistence length ( $\ell_P$ , 47 nm).<sup>17</sup>

It is easy to imagine that the collective deformations of a uniformly buckled rod, when subjected to extreme bending, could be superseded by localized excitations that render short stretches of the chain very pliable. Recent theoretical work has sought to account in this manner for the anomalous rates of Cloutier and Widom by positing a wormlike chain model incorporating such thermally-excited ‘melts’ or ‘kinks’. The apparent persistence length (after integrating over the excitations) of such a model is greater than that of the conventional wormlike chain: melting influences bending. In this work we demonstrate that, conversely, it follows that bending influences melting: the apparent free energy of forming such an excitation (after integrating over the spatial degrees of freedom) is less than the ‘bare’ free energy directly entering into the model Hamiltonian. In §IV we show that as originally parameterized, the melttable wormlike chain theory of Yan and Marko<sup>19,20</sup> improperly accounts for this entropic stabilization of enhanced-flexibility excitations. Specifically, with the original thermodynamic penalty for local melting, that model appears to be inconsistent with well-characterized thermodynamics of bulk DNA melting. We discuss this double-counting of certain entropic gains due to melting, and in §V propose a revised parameterization for the ‘melttable’ WLC model. In §VII we outline a new method for eval-

uating  $J$ -factors that characterize cyclization kinetics, from enhanced-sampling Monte Carlo (MC) simulations. This method could be applied to any comparably coarse-grained model and, unlike existing methods, incurs similar computational cost for both long and short DNA chains. In §VIII we explore cyclization rates using these MC simulations, finding that such excitations, properly parameterized by well-established DNA melting thermodynamics, produce no practical difference in cyclization rates for DNA chains longer than  $\sim 120$  bp. Only for shorter chains do models including melting excitations produce significantly higher cyclization rates than the unmelttable WLC model, though this rate enhancement is lower than that seen in some experiments.

## II. THE WORMLIKE CHAIN MODEL WITH THERMAL EXCITATIONS OF ENHANCED FLEXIBILITY

The WLC model envisions DNA as a single chain of uniform-length segments (each representing a fixed number of basepairs) connected at a set of nodes, where the bend angle at a given node  $i$  is allowed to fluctuate according to a bending energy quadratic in the local curvature

$$\beta E_i^{\text{WLC}} = \frac{1}{2} \frac{\ell_P}{d} |\hat{t}_{i+1} - \hat{t}_i|^2 \quad (1a)$$

$$= \frac{\ell_P}{d} (1 - \cos \theta_i). \quad (1b)$$

Here  $\beta = (k_B T)^{-1}$ ,  $\ell_P$  is the persistence length,  $\hat{t}_i$  is the unit-length vector pointing from node  $i - 1$  to node  $i$ ,  $d$  is the contour length separating adjacent nodes, and  $\theta_i = \cos^{-1}(\hat{t}_{i+1} \cdot \hat{t}_i)$  is the local bending angle (Fig. 1a).

The melttable WLC (due to Yan and Marko<sup>19,20</sup>) adds an additional set of variables, the hybridization states of those nodes: each node has two internal states, notionally corresponding to hybridized and melted local structures. Fluctuations in base pairing in this model are represented by changes in the hybridization state. The free energy associated with a single node in hybridization state  $m$  and bent at an angle  $\theta$  is

$$\beta E^{\text{MWLC}}(m, \theta) = \delta_{m,0} \left[ \frac{\ell_P^{\text{H}}}{d} (1 - \cos \theta) \right] + \delta_{m,1} \left[ \beta \Delta \mu(T) + \frac{\ell_P^{\text{M}}}{d} (1 - \cos \theta) \right]. \quad (2)$$

where  $\ell_P^{\text{H}}$  is the persistence length of hybridized DNA and  $\ell_P^{\text{M}}$  is the persistence length of molten DNA. We have defined  $m = 0$  as the hybridized state and  $m = 1$  as the molten state. The Kronecker delta,  $\delta_{i,j}$ , is 0 if  $i \neq j$  and 1 if  $i = j$ . The thermodynamic penalty for melting  $\Delta \mu$  is the reversible work required to disrupt base pairing at a node of the chain that has a fixed bend angle  $\theta = 0$ .

We also compare the MWLC with two other models originally formulated from similar perspectives. The kinkable wormlike chain (KWLC)<sup>21</sup> posits a completely

flexible molten state,  $\ell_P^M = 0$ . The spontaneous bend model maintains some rigidity in the molten state but introduces a preferred bend angle  $\cos^{-1} \gamma$ :

$$\beta E^{\text{SB}}(m, \theta) = \delta_{m,0} \left[ \frac{\ell_P^H}{d} (1 - \cos \theta) \right] + \delta_{m,1} \left[ \beta \Delta \mu(T) + \frac{\ell_P^M}{d} (\cos \theta - \gamma)^2 \right]. \quad (3)$$

Nevertheless, the specification of a preferred bend angle seems less physically plausible for bare DNA than for protein-bound DNA, and the assumption of completely flexible molten regions is unnecessarily extreme for our purposes. Thus we focus primarily on the role of thermally-excited local regions of enhanced flexibility ('melts') in the cyclization kinetics of DNA.

In these models a chain of contour length  $(N+1)d$  has  $N+2$  nodes (which we number 0 to  $N+1$ ) connected by  $N+1$  fixed-length links, and  $N$  bending degrees of freedom (Fig. 1a). Lacking bending energy, the first and last nodes can be ignored in writing the chain's total energy:

$$E_{\text{chain}}[\{m_i\}, \{\theta_i\}] = \sum_{i=1}^N E(m_i, \cos \theta_i). \quad (4)$$

Because nodes contribute to  $E_{\text{chain}}$  in an additive fashion, they fluctuate in a statistically independent manner (in the absence of collective constraints).

In considering a model chain that is inextensible along its contour, we neglect potential effects of stretching fluctuations on cyclization. This approximation is supported by estimates of the dsDNA stretch modulus,  $\sim 1000$  pN,<sup>2,22</sup> and by the threshold force,  $\sim 65$  pN, for the overstretching transition to a form  $\sim 1.7\times$  longer.<sup>3</sup> Both values well exceed the typical range of forces anticipated in the course of loop closure. This model only implicitly accounts for structural aspects on length scales smaller than a single link, such as the detailed atomic structure of the DNA basepair.

### III. CYCLIZATION KINETICS

For decades, cyclization reactions have been a popular experimental tool for measuring DNA bending elasticity.<sup>4</sup> A population of DNA chains of given length and sequence, with 'sticky' single-stranded overhangs on each end, are incubated in solution with a DNA ligase, which can covalently link the two ends together when they come into spatial proximity and appropriate relative orientation. Samples are taken at periodic time intervals and run on a gel to quantify the fluorescence from bands corresponding to populations of circularized monomers, linear dimers, and so on. Dividing the rate constant for cyclization  $k_{\text{cyc}} \equiv \lim_{t \rightarrow 0} [dC(t)/dt] / M_0$  by the rate constant

for dimerization  $k_{\text{dim}} \equiv \lim_{t \rightarrow 0} [dD(t)/dt] / M_0^2$ , controls in certain limits for the ligase enzyme's propensity to ligate apposed sticky ends, yielding a measure of chain flexibility, the so-called  $J$ -factor,  $J = 2k_{\text{cyc}}/k_{\text{dim}}$ . Here  $C(t)$  and  $D(t)$  are the concentrations of ligated circular monomers and ligated linear dimers, respectively, and  $M_0$  is the initial concentration of unligated monomers.

In the case of low ligase concentration, ligation should proceed at a rate proportional to the equilibrium population of precursors adopting ligation-competent conformations. We define a monomer conformation as ligation-competent when one end lies within a critical volume  $v^*$  (of critical radius  $R^*$ ) of the other end of the same molecule, while the relative polar and azimuthal angles of closure ( $\theta_N - \theta_0$  and  $\phi_N - \phi_0$ , respectively) lie below small threshold values ( $\theta^*$  and  $\phi^*$ , respectively). Therefore at short times the concentration of ligated circular monomers will increase with rate

$$\frac{dC(t)}{dt} = k_0 M_0 \times P(|\mathbf{R}_N - \mathbf{R}_0| < R^*, \cos(\theta_N - \theta_0) > \cos \theta^*, |\phi_N - \phi_0| < \phi^*) \quad (5)$$

Here  $k_0$  is the ligation rate for ligation-competent monomers, and  $P(|\mathbf{R}_N - \mathbf{R}_0| < R^*, \cos(\theta_N - \theta_0) > \cos \theta^*, |\phi_N - \phi_0| < \phi^*)$  is the probability that the chain is considered ligation-competent. The concentration of ligated linear dimers will correspondingly increase as

$$\frac{dD(t)}{dt} = 2k_0 V M_0^2 \times \overline{P}(|\mathbf{R}_N - \mathbf{R}_0| < R^*, \cos(\theta_N - \theta_0) > \cos \theta^*, |\phi_N - \phi_0| < \phi^*) \quad (6)$$

where

$$\overline{P}(|\mathbf{R}_N - \mathbf{R}_0| < R^*, \cos(\theta_N - \theta_0) > \cos \theta^*, |\phi_N - \phi_0| < \phi^*) = \frac{v^*}{V} \left( \frac{1 - \cos \theta^*}{2} \right) \frac{\phi^*}{2\pi} \quad (7)$$

is the probability of one end of a given molecule lying within the critical volume of the end of another molecule, within the critical relative polar and azimuthal angles. The restriction enzymes used to generate sticky ends in the experiments discussed here (*EagI* and *HindIII*) operate on reverse-palindromic sequences, so each monomer has an identical overhanging sequence on each end, and thus either end can be ligated to another monomer, producing the factor of 2 in Eq. (6).<sup>24</sup>

We assume that end-to-end distances at which ligation occurs are smaller than any length scale characterizing chain statistics, and thus that the probability density is uniform and isotropic within the critical volume and critical angles,

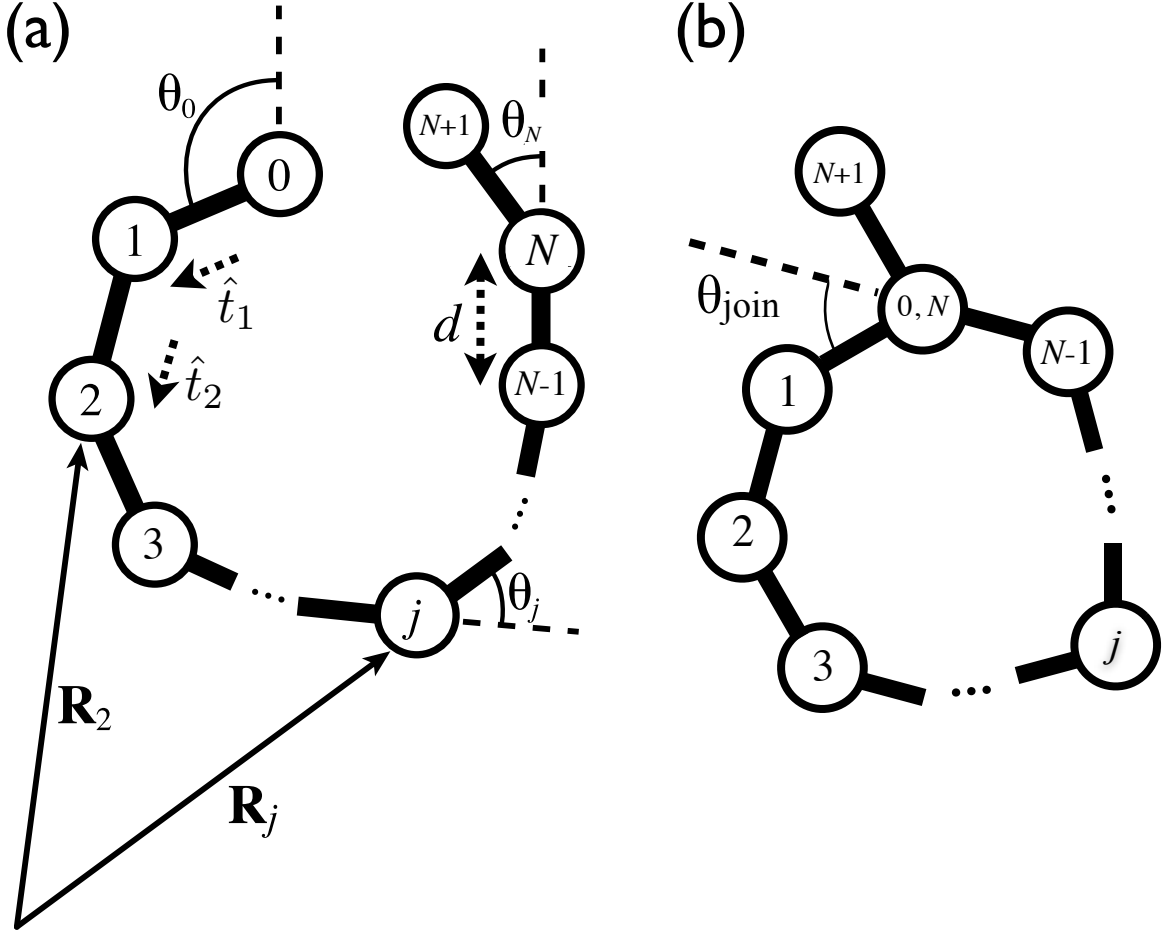


FIG. 1. Geometry of a melttable WLC comprising  $N + 1$  links. (a) Chain cyclization amounts to imposing the constraints  $\mathbf{R}_0 = \mathbf{R}_N$ ,  $\theta_0 = \theta_N$ , and  $\phi_0 = \phi_N$ ; where  $\phi_0$  and  $\phi_N$  are azimuthal angles (not shown) corresponding to the polar angles  $\theta_0$  and  $\theta_N$ , respectively. Note that the angle  $\theta_0$  is defined with reference to the vector  $\mathbf{R}_N - \mathbf{R}_{N-1}$ , which is oriented vertically in this example configuration. (b) To determine the propensity of apposed chain ends to attain similar relative orientation, biased simulations (see §VI) calculate the conditional probability distribution of  $\cos \theta_{\text{join}}$ , given  $\mathbf{R}_N = \mathbf{R}_0$ , where  $\theta_{\text{join}}$  is the angle between the vectors  $\mathbf{R}_N - \mathbf{R}_{N-1}$  and  $\mathbf{R}_1 - \mathbf{R}_0$ .

$$P(|\mathbf{R}_N - \mathbf{R}_0| < R^*, \cos(\theta_N - \theta_0) > \cos \theta^*, |\phi_N - \phi_0| < \phi^*) = v^*(1 - \cos \theta^*) \phi^* \langle \delta(\mathbf{R}_N - \mathbf{R}_0) \delta(1 - \cos[\theta_N - \theta_0]) \delta(\phi_N - \phi_0) \rangle_N \quad (8)$$

where the angled brackets with subscript  $N$  denote a canonical average over the  $2N$  angular degrees of freedom (two at each internal node) of an  $(N+2)$ -node chain according to the specified energy function (in our case  $E_{\text{chain}}$  in Eq. (4)). The Dirac  $\delta(x)$  is 0 when  $x \neq 0$  and integrates to unity over any region including  $x = 0$ , thus the latter two  $\delta$ -functions vanish unless the tangent vectors  $(\mathbf{R}_1 - \mathbf{R}_0)/d$  and  $(\mathbf{R}_{N+1} - \mathbf{R}_N)/d$  are identical. (See Fig. 1a.)

Under these conditions, the  $J$ -factor reduces to an effective probability density of the unligated chain forming a transient closed loop, with zero relative polar and az-

imuthal angles:

$$J = 2 \frac{k_{\text{cyc}}}{k_{\text{dim}}} \quad (9a)$$

$$= 2M_0 \lim_{t \rightarrow 0} \left[ \frac{dC(t)}{dt} / \frac{dD(t)}{dt} \right] \quad (9b)$$

$$= \frac{4\pi}{v^*(1 - \cos \theta^*) \phi^*} \times P(|\mathbf{R}_N - \mathbf{R}_0| < R^*, \cos(\theta_N - \theta_0) > \cos \theta^*, |\phi_N - \phi_0| < \phi^*) \quad (9c)$$

$$= 4\pi \langle \delta(\mathbf{R}_N - \mathbf{R}_0) \delta(1 - \cos[\theta_N - \theta_0]) \delta(\phi_N - \phi_0) \rangle_N. \quad (9d)$$

In §VII we compute the probability density  $\langle \delta(\mathbf{R}_N -$

$\mathbf{R}_0) \delta(1 - \cos[\theta_N - \theta_0]) \delta(\phi_N - \phi_0) \rangle_N$  for the melttable WLC model outlined in §II.

When ligase concentration is not sufficiently low, the ligation rate of ligation-competent conformations becomes comparable to the rate of their formation from ligation-incompetent precursors. In this concentration regime ligation rates will thus depend on the dynamics of formation of ligation-competent conformations, not just on their equilibrium probabilities. For circular monomers, this reflects the rate of motion in the coordinate of the end-to-end distance. For linear dimers, kinetics will be diffusion-controlled. Thus the ratio of circularization and dimerization rates no longer simply reflects an equilibrium probability, and the  $J$ -factor no longer reflects equilibrium free energies. Du, *et al.* showed that at 21°C the threshold ligase concentration was  $\sim 100$  units/ml, and thus argued that Cloutier and Widom's ligase concentrations of 150-250 units/ml were sufficiently high that they did not measure equilibrium bending propensities. Recent experiments by Forties, *et al.*<sup>25</sup> complicate the picture by indicating a threshold at 37°C of  $\sim 400$  units/ml, suggesting that Cloutier and Widom's experiments at 30°C might have near-threshold ligase concentrations. See also Peters and Maher's review for further discussion.<sup>26</sup>

In the limit of high ligase concentration, the rate-limiting step is formation of a ligation-competent conformation, and hence the  $J$ -factor would reflect the effective reaction rate of aligning the ends. This quantity can in principle be calculated from molecular dynamics simulations of models with more detailed representation of the DNA: though the relevant dynamical modes might be complex, a first approximation might assume diffusive motion in the end-to-end distance and join angle, and subsequently solve for the diffusion-controlled steady state using a Langevin or Fokker-Planck approach. Comparison with experiments at high ligase concentration could thus yield insights into DNA dynamics on these short length scales. Cheng, *et al.* recently conducted similar experimental and computational examination of the cyclization dynamics of single-stranded DNA.<sup>27</sup> In the rest of this paper, we restrict our discussion to the regime of low ligase concentration.

#### IV. CONNECTION WITH LATTICE MODELS

One-dimensional lattice models for fluctuations in DNA hybridization have a long history in the study of bulk melting thermodynamics.<sup>28</sup> Systems of interest are typically free of any external mechanical constraints, so the models typically make no statement about chain structure. In particular they assign free energy to a short segment of the chain based only on its hybridization state  $m$ :

$$\overline{E}(m) = \delta_{m,1} \overline{\Delta\mu}(T). \quad (10)$$

The chain's total free energy in this description similarly depends only on the hybridization variables,  $\overline{E}_{\text{chain}}[\{m_i\}] = \sum_i \overline{E}(m_i)$ . Implicit in the free energy assignment of Eq. (10) is an integration over conformational fluctuations weighted by the Boltzmann distribution,

$$\begin{aligned} & \exp(-\beta \overline{E}_{\text{chain}}[\{n_i\}]) \\ &= \int_{-1}^1 d(\cos \theta_1) \int_{-1}^1 d(\cos \theta_2) \dots \int_{-1}^1 d(\cos \theta_{N-1}) \\ & \quad \times \exp(-\beta E_{\text{chain}}[\{n_i\}, \{\theta_i\}]). \end{aligned} \quad (11)$$

Analytical evaluation of this integral is straightforward for the melttable WLC, since the Boltzmann weight factorizes.

The partition function at a given node for hybridization state  $m$  (integrating Eq. (2) over all bend angles  $\cos \theta$ ) is

$$q(m) = \int_{-1}^1 d(\cos \theta) e^{-\beta E(m, \cos \theta)} \quad (12a)$$

$$\begin{aligned} &= \frac{d}{\ell_P^H + \delta_{m,1} (\ell_P^M - \ell_P^H)} \exp\{-\delta_{m,1} \beta \Delta\mu\} \\ & \quad \times \left(1 - \exp\left\{-\frac{2}{d} [\ell_P^H + \delta_{m,1} (\ell_P^M - \ell_P^H)]\right\}\right). \end{aligned} \quad (12b)$$

The free energy difference between melted ( $m = 1$ ) and non-melted ( $m = 0$ ) states is  $\Delta F = F(1) - F(0) = -k_B T \ln[q(1)/q(0)]$ . We arrive at a relationship between the 'bare' free energy of melting ( $\Delta\mu^{\text{MWLC}}$ ) and its 'renormalized' counterpart ( $\overline{\Delta\mu}$ ) that accounts for the influence of bending fluctuations:

$$\beta \overline{\Delta\mu} = \beta \Delta\mu^{\text{MWLC}} - \ln \left[ \frac{\ell_P^H}{\ell_P^M} \left( \frac{1 - \exp(-2\ell_P^M/d)}{1 - \exp(-2\ell_P^H/d)} \right) \right] \quad (13a)$$

$$\approx \beta \Delta\mu^{\text{MWLC}} - \ln \left( \frac{\ell_P^H}{\ell_P^M} \right). \quad (13b)$$

The second term in the RHS of Eq. (13a) accounts for the greater range of bending motion available to a molten segment of the chain, which is considerable. For the length parameters employed by Yan and Marko<sup>19</sup> ( $\ell_P^H = 50$  nm,  $\ell_P^M = 1$  nm,  $d = 1$  nm), the renormalization evaluates to

$$\overline{\Delta\mu} - \Delta\mu^{\text{MWLC}} \approx -3.8 k_B T. \quad (14)$$

Similar analysis for the KWLC yields

$$\beta \overline{\Delta\mu} = \beta \Delta\mu^{\text{KWLC}} - \ln \left( \frac{2 \ell_P^H}{d[1 - \exp(-2 \ell_P^H/d)]} \right) \quad (15)$$

$$\approx \beta \Delta\mu^{\text{KWLC}} - \ln \left( \frac{2 \ell_P^H}{d} \right) \quad (16)$$

$$\approx \beta \Delta\mu^{\text{KWLC}} - 4.6, \quad (17)$$

and for the spontaneous bend model

$$\beta\overline{\Delta\mu} = \beta\Delta\mu^{\text{SB}} - \ln \sqrt{\frac{\pi}{2d\ell_P^{\text{M}}}} \frac{\ell_P^{\text{H}}}{[1 - \exp(-2\ell_P^{\text{H}}/d)]} \quad (18)$$

$$\times \left\{ \text{erf} \left[ \sqrt{\frac{\ell_P^{\text{M}}}{2d}}(\gamma + 1) \right] - \text{erf} \left[ \sqrt{\frac{\ell_P^{\text{M}}}{2d}}(\gamma - 1) \right] \right\}. \quad (19)$$

Conversely, integrating over melt degrees of freedom renormalizes the persistence length.<sup>29,30</sup>

## V. PARAMETERIZATION

The large body of work on DNA melting provides good estimates for the thermodynamic parameters appropriate to one-dimensional lattice models such as Eq. (10). Bulk melting free energies include a renormalization [Eq. (13)] due to integrating out bending fluctuations, therefore proper use of melting data to parameterize microscopic models must involve disentangling this renormalization. If each node or lattice site represents three basepairs, then we expect  $\overline{\Delta\mu} \approx 8 - 11 k_B T$  at ambient conditions.<sup>31</sup> Yan and Marko presented an argument for this range of values,<sup>19</sup> but no distinction was made there between  $\Delta\mu^{\text{MWLC}}$  and  $\overline{\Delta\mu}$ . If we assign the inferred melting penalty to the bare free energy rather than the renormalized value (*i.e.*,  $\Delta\mu^{\text{MWLC}} \approx 8 - 11 k_B T$  and  $\overline{\Delta\mu} \approx 4 - 7 k_B T$ ), hybridization statistics of the meltable WLC at room temperature are not consistent with experiment. Specifically, a three-base region is melted with a probability between  $7.2 \times 10^{-4}$  and  $1.4 \times 10^{-2}$  (*i.e.*, between  $e^{-4.2}$  and  $e^{-7.2}$ ), instead of between  $1.7 \times 10^{-5}$  and  $3.3 \times 10^{-4}$  (*i.e.*, between  $e^{-8}$  and  $e^{-11}$ ) as suggested by thermodynamic measurements.

Furthermore, this significantly-increased probability of melts yields chains that are too compliant to be consistent with longer length scale force-extension experiments. The persistence length is determined by  $\langle \cos \theta \rangle$ , the average cosine of the angle at a given node. For a meltable node this average cosine is an equilibrium-weighted linear combination of the average cosines of a melted or hybridized node,

$$\langle \cos \theta \rangle = \frac{e^{-\beta\overline{\Delta\mu}}}{1 + e^{-\beta\overline{\Delta\mu}}} \langle \cos \theta \rangle_{\text{M}} + \frac{1}{1 + e^{-\beta\overline{\Delta\mu}}} \langle \cos \theta \rangle_{\text{H}}, \quad (20)$$

producing for the MWLC an apparent persistence length (averaged over melt fluctuations)

$$\ell_P^{\text{MWLC}} = \frac{d}{1 - \langle \cos \theta \rangle^{\text{MWLC}}} = \frac{1 + e^{-\beta\overline{\Delta\mu}}}{(e^{-\beta\overline{\Delta\mu}}/\ell_P^{\text{M}}) + (1/\ell_P^{\text{H}})}. \quad (21)$$

In particular, for the range of bare melt free energies  $\Delta\mu^{\text{MWLC}} = 8 - 11 k_B T$ , the persistence length is renormalized to  $\ell_P^{\text{MWLC}} \approx 30\text{-}48$  nm.

To maintain a given apparent persistence length  $\ell_P$  for a given renormalized melt free energy  $\beta\overline{\Delta\mu}$ , the hybridized persistence length  $\ell_P^{\text{H}}$  is thus a function of the melt persistence length  $\ell_P^{\text{M}}$ :

$$\ell_P^{\text{H,MWLC}} = \frac{\ell_P}{1 - \left( \frac{\ell_P}{\ell_P^{\text{M}}} - 1 \right) e^{-\beta\overline{\Delta\mu}}}. \quad (22)$$

For the KWLC model we find

$$\ell_P^{\text{H,KWLC}} = \frac{\ell_P}{1 - \left( \frac{\ell_P}{d} - 1 \right) e^{-\beta\overline{\Delta\mu}}}, \quad (23)$$

and for the spontaneous bend model with at least modest melt persistence length

$$\ell_P^{\text{H,SB}} \approx \frac{\ell_P}{1 - \left( \frac{\ell_P}{d} [\gamma - 1] - 1 \right) e^{-\beta\overline{\Delta\mu}}}. \quad (24)$$

Based on Eqs. (14) and (22) we determine a more realistic value for the bare free energy of local melting in the MWLC,  $\Delta\mu^{\text{MWLC}} \approx 12 - 15 k_B T$ , which is consistent with bulk thermodynamics by construction. The increased  $\Delta\mu^{\text{MWLC}}$ , designed to avoid double-counting entropic gains due to increased flexibility, will clearly act to suppress melting and to diminish its importance in the kinetics of cyclization. We have quantified this suppression using computer simulations.

## VI. RELATING $J$ -FACTORS TO DISTRIBUTIONS OF END-TO-END DISTANCE AND ANGLE IN A DISCRETIZED CHAIN

The  $J$ -factor was expressed in §III in terms of an equilibrium probability density that lends itself to straightforward evaluation in computer simulations. For a chain of contour length  $(N + 1)d$ , whose  $N + 2$  nodes are located at positions  $\{\mathbf{R}_0, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_{N+1}\}$ , we rewrite Eq. (9d) as

$$J(N) = 4\pi \frac{Q_N^{(\text{loop})}}{Q_N^{(0)}}, \quad (25)$$

We have defined partition functions  $Q_N^{(\text{loop})}$  for a cyclized  $(N + 2)$ -node meltable WLC, and  $Q_N^{(0)} = q^N$  for the same molecule absent end constraints. The single-node partition function  $q$  can be evaluated as

$$q = 2\pi \int_{-1}^1 d(\cos \theta) e^{-\beta\tilde{E}(\cos \theta)}, \quad (26)$$

for apparent bending energy

$$\tilde{E}(\cos \theta) = -k_B T \ln \sum_{n=0}^1 e^{-\beta E(n, \cos \theta)}. \quad (27)$$

Thus,

$$J(N) = 4\pi \int d(\cos \theta_1) \cdots \int d(\cos \theta_N) \quad (28)$$

$$\times \int d\phi_1 \cdots \int d\phi_N \frac{e^{-\beta \sum_{j=1}^N \tilde{E}(\cos \theta_j)}}{q^N}$$

$$\times \delta(\mathbf{R}_N - \mathbf{R}_0) \delta(1 - \cos[\theta_N - \theta_0]) \delta(\phi_N - \phi_0)$$

The first  $\delta$ -function imposes no constraint on the variables  $\cos \theta_N$  or  $\phi_N$  (see Fig. 1a), and thus integrating over both is trivial, giving

$$J(N) = 4\pi q^{-1} \left\langle \delta(\mathbf{R}_N - \mathbf{R}_0) e^{-\beta \tilde{E}(\cos \theta_0)} \right\rangle_{N-1}. \quad (29)$$

We then use the delta-function identity

$$e^{-\beta \tilde{E}(\cos \theta_0)} = \quad (30)$$

$$\int d(\cos \theta_{\text{join}}) e^{-\beta \tilde{E}(\cos \theta_{\text{join}})} \delta(\cos \theta_{\text{join}} - \cos \theta_0),$$

where the join angle  $\theta_{\text{join}}$  describes the relative orientations of the chain's ends,  $\cos \theta_{\text{join}} = (\mathbf{R}_N - \mathbf{R}_{N-1})(\mathbf{R}_1 - \mathbf{R}_0)/d^2$ , when it is poised for cyclization,  $\mathbf{R}_N = \mathbf{R}_0$  (i.e.,  $\mathbf{R} = \mathbf{0}$ ). (See Fig. 1b.) Finally, we multiply and divide by  $\langle \delta(\mathbf{R}_N - \mathbf{R}_0) \rangle_{N-1}$ , substitute the probability density at zero end-to-end extension,  $p_N(\mathbf{R}_0 = \mathbf{R}_N) = \langle \delta(\mathbf{R}_N - \mathbf{R}_0) \rangle_{N-1}$ , and introduce the conditional probability density at a given bend angle of a particular node (given a cyclized chain),

$$p_N(\cos \theta_{\text{join}} | \mathbf{R} = \mathbf{0}) = \quad (31)$$

$$\frac{\langle \delta(\mathbf{R}_N - \mathbf{R}_0) \delta(\cos \theta_{\text{join}} - \cos \theta_0) \rangle_{N-1}}{\langle \delta(\mathbf{R}_N - \mathbf{R}_0) \rangle_{N-1}}.$$

These manipulations enable us to express the  $J$ -factor in terms of quantities readily estimated from sampling of canonically distributed microstates,

$$J(N) = 4\pi q^{-1} p_N(\mathbf{R} = \mathbf{0}) \int d(\cos \theta_{\text{join}}) \quad (32)$$

$$\times e^{-\beta \tilde{E}(\cos \theta_{\text{join}})} p_N(\cos \theta_{\text{join}} | \mathbf{R} = \mathbf{0}).$$

## VII. MONTE CARLO SIMULATIONS

Eq. (25), together with an expansion of the Boltzmann distribution in spherical harmonics, was used by Yan and Marko<sup>19</sup> to evaluate  $J(N)$  for the melttable WLC. Their approach becomes quite numerically demanding, however, for the short chains of interest here. We have instead used Monte Carlo simulations to estimate the probability distributions appearing in Eq. (32). This approach also permits the simple extension of the model to more complicated energetics, involving considerations such as volume exclusion and DNA helicity.<sup>30</sup>

We employ two kinds of trial moves to efficiently navigate conformational space while maintaining the geometric constraints of the molecule. ‘Free’ rotations divide the chain in two by selecting a randomly chosen node, then rotate the two segments relative to one another about a randomly chosen axis through the dividing node. Crankshaft moves rotate a segment of the chain between two randomly chosen nodes, about the axis connecting them. The number of contiguous nodes rotated in crankshaft moves is chosen from a uniform distribution on the interval [1,10]. Angles for all rotations are selected at random from Gaussian distributions whose variances were adjusted until roughly 1/2 of all trial moves are accepted. Melt moves attempt to change the hybridization state of a randomly chosen basepair, either from hybridized to melted or *vice versa*.

Calculating cyclization probabilities for short DNA chains via standard equilibrium sampling is computationally challenging because of the extreme energetic cost of juxtaposing the two ends with the correct orientation. We estimated the equilibrium average on the right-hand side of Eq. (25) by sampling from the Boltzmann distribution using Wang-Landau sampling,<sup>32</sup> an adaptive method for broadly sampling fluctuations of a chosen order parameter without any prior knowledge about the details or even the gross shape of the probability distribution.

This biased sampling approach allowed us to construct the distribution  $p_N(\mathbf{R}) = \langle \delta(\mathbf{R} - (\mathbf{R}_N - \mathbf{R}_0)) \rangle_N$  of the end-to-end vector  $\mathbf{R}$  for an uncyclized chain, over the full range of  $|\mathbf{R}|$  (Fig. 2). Similarly biased sampling, this time of a cyclized chain with no bending energy at one node, was used to construct the conditional angular distribution at that freely-jointed node (Fig. 3):

$$p_N(\cos \theta_{\text{join}} | \mathbf{R} = \mathbf{0}) = \frac{\langle \delta(\cos \theta_{\text{join}} - d^{-2}[\mathbf{R}_1 - \mathbf{R}_0] \cdot [\mathbf{R}_N - \mathbf{R}_{N-1}]) \delta(\mathbf{R}_N - \mathbf{R}_0) \rangle_N}{\langle \delta(\mathbf{R}_N - \mathbf{R}_0) \rangle_N}. \quad (33)$$

For each chain length and melting free energy, 3 different sampling trajectories were run from identical initial conditions but with different random number sequences to calculate uncertainties. Each Wang-Landau iteration ran

until each of the 1000 spatial bins (or 200 angular bins) was visited at least 100 times. The 10 spatial bins at greatest extension were excluded from this criterion due to their exceedingly small probabilities. The increment

to the Wang-Landau bias began at  $1 k_B T$  and decreased by 10% at each subsequent iteration, until the increment reached  $10^{-7} k_B T$  ( $10^{-8} k_B T$  for angular simulations of chains shorter than 120 bp). A few angular simulations of shorter chains achieved greater precision with a 20% (rather than 10%) decrease in increment per iteration.

Probability distributions of the end-to-end distance  $\mathbf{R}$  and join angle  $\theta_{\text{join}}$ , for meltable wormlike chains of varying length and melt free energy, are shown in Figs. 2 and 3. For meltable chains, the end-to-end distribution approximates at large distances that of the un-meltable chain. As end-to-end distance decreases below the range of typical fluctuations of a WLC, melting begins to significantly facilitate achieving loop closure of sufficiently short chains. Their significant energetic cost requires that substantial bending forces develop before such local excitations become favorable.

The general shape of  $p_N(\mathbf{R})$ , which has been discussed at length elsewhere,<sup>33</sup> can be simply understood in terms of entropic elasticity at large  $\mathbf{R}$  and bending energy at small  $\mathbf{R}$ . Straightening a chain imposes high entropic cost, so that probability decreases sharply with  $\mathbf{R}$  at large extension. Imposing a short end-to-end distance requires significant bending and thus high enthalpic cost, explaining the low probability for very bent chains. The longest chain is easiest to cyclize since the total bend can be distributed over the largest number of basepairs.

At the very shortest end-to-end distances, meltable chains show a slight increase in  $p_N(\mathbf{R})$  as distance decreases, for geometric reasons. Imposing such a tight bend essentially requires a melt at the middle node of the chain; once that node is melted it serves as a nearly free joint between two nearly rigid chain segments on either side. Taking the joint as the origin of a molecular reference frame, the two endpoints thus effectively explore the surface of a sphere. As they approach one another in this two-dimensional geometry, the number of accessible conformations  $\Omega$  with a given value of  $R$  grows only linearly,  $\Omega \sim R$  (rather than quadratically, as it would for a very flexible chain). The probability per unit volume  $p_N(\mathbf{R})$  thus scales here roughly as  $\Omega(R)/(4\pi R^2) \sim 1/R$ . Because the chain ends are not strictly confined to a two-dimensional surface,  $p_N(\mathbf{R})$  does not diverge as  $R \rightarrow 0$ , but it does grow over a small range of  $R$ . A similar scenario, now involving melts at nodes one or two removed from the central one, accounts for the small oscillations at slightly longer end separations. Consistent with this explanation, the local maxima in probability appear at end-to-end distances that are even integer multiples of the link length.

For unmeltable chains, the join angle statistics (for a chain that meets at its ends but lacks bending energy at that meeting point) in Fig. 3 show a probability maximum at intermediate angles. This stems from the enthalpic preference for equally distributing the bend over the entire contour of the chain, producing a teardrop-shaped chain contour with join angle intermediate between  $0^\circ$  (straight) and  $180^\circ$  (fully bent).<sup>34</sup> Meltable short

chains differ substantially. In accommodating loop closure, the energetic cost of severe bending distributed across the entire uniformly curved chain (without any melts) is substantially higher than that of a single melt at the middle basepair opposite the junction. The latter state is energetically minimized with a join angle of  $\sim 180^\circ$  and very little bending on either side. Conversely, the low probability of a join angle of  $0^\circ$  ( $\cos \theta_{\text{join}} = 1$ ) reflects the need for formation of a second melt, essentially required to bring the ends of such short chains into common orientation.

## VIII. $J$ -FACTORS

We have confirmed numerically that this MC approach is consistent with the spherical harmonic expansion used by Yan and Marko.<sup>19</sup> For the chains we compared, of length 135-225 bp, the deviations between their results and our calculations are not statistically significant (data not shown). Henceforth we calculate MWLC  $J$ -factors for parameters that differ in two respects from those of Yan and Marko<sup>19</sup>: (i) melted sections are slightly stiffer,  $\ell_P^M = 2.5$  nm, than in their calculations (where  $\ell_P^M = 1$  nm); and (ii) the range of melting free energy values we explore extends to the stiffer thermodynamic penalties suggested in §V. We argue that the longer persistence length of unhybridized chains is more consistent with experimental estimates for single-stranded DNA (ssDNA) bending elasticity,<sup>2,35</sup> since a melted section composed of two single strands of DNA in close apposition would be expected to have a persistence length at least twice that of ssDNA.

All of the model results we report, like those of Yan and Marko, neglect the helical nature of hybridized dsDNA, implicitly assuming that two ends juxtaposed with any helical phasing are ligation-competent. The calculations thus represent an upper bound for the  $J$ -factor of a helical wormlike chain polymer, a bound that is tightest for chains that are an integral number of helical repeats in length.

Fig. 4 shows our numerical results for contour lengths of 10-120 nm, with melting energies from Yan and Marko<sup>19</sup> (a) or with values of  $\Delta\mu$  that are consistent with bulk melting thermodynamics (b). The shift from  $\Delta\mu = 8 - 11 k_B T$  (Fig. 4a) to  $\overline{\Delta\mu} = 8 - 11 k_B T$  (Fig. 4b) represents an increase in melting free energy of  $3.7 k_B T$  and hence considerably suppresses melts. In fact, whereas for  $\Delta\mu = 8 - 11 k_B T$  only the 360 bp chain  $J$ -factors are unaffected by melts, for  $\overline{\Delta\mu} = 8 - 11 k_B T$  ( $\Delta\mu = 11.7 - 14.7 k_B T$ ) meltable chains as short as 150 bp cyclize with propensities indistinguishable from the melt-free limit,  $\overline{\Delta\mu} = \Delta\mu = \infty$ . For shorter chains the meltable models produce a  $J$ -factor significantly greater than that for the melt-free ( $\Delta\mu = \overline{\Delta\mu} = \infty$ ) model, because cyclization dictates a bend so sharp that the energetic gain from localizing the bend at a floppy melted section overcomes the energetic cost of forming the melt.



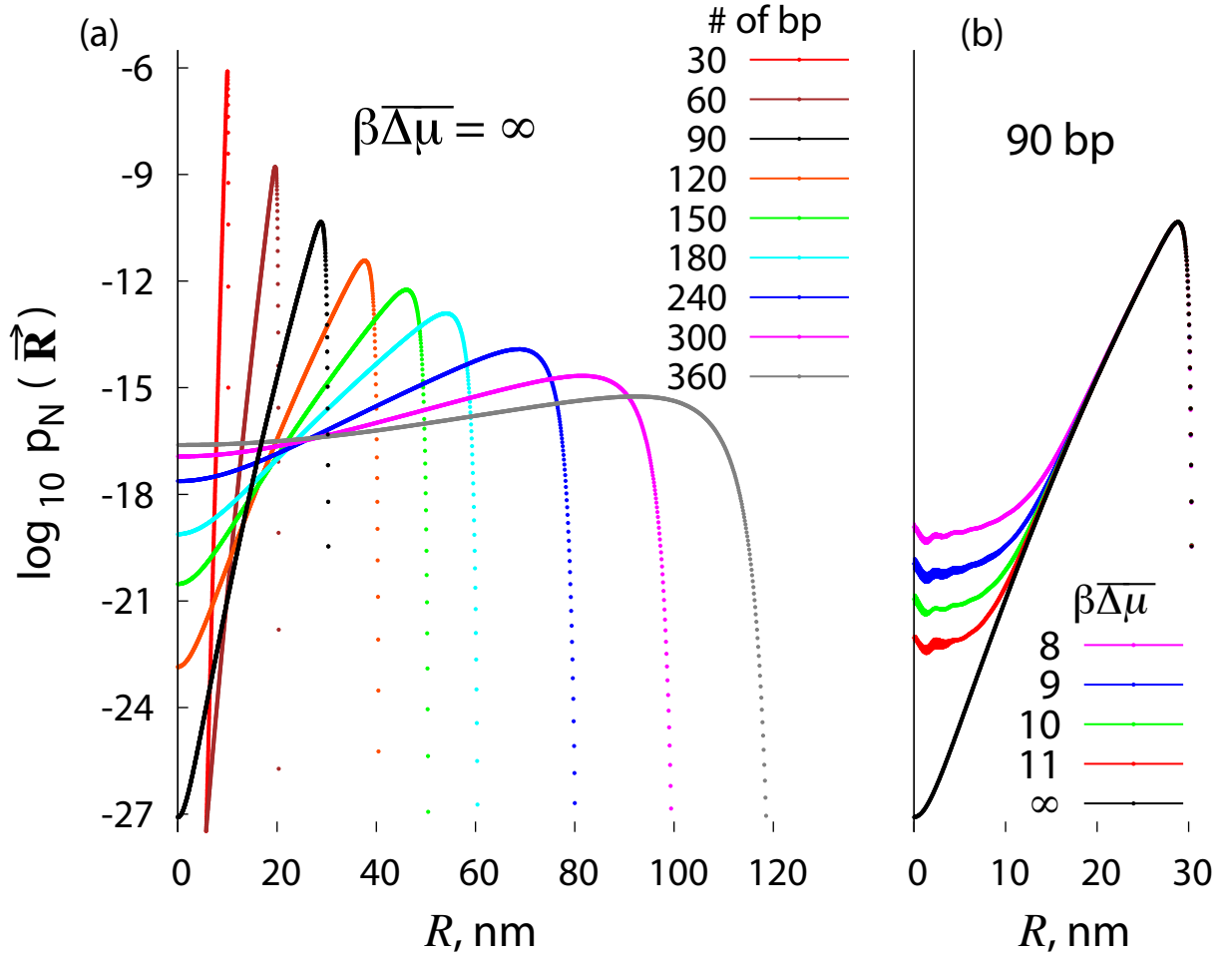


FIG. 2. Bending statistics for meltable wormlike chains of varying contour length and melting thermodynamics. Error bars show  $\pm 2$  standard errors of the mean (s.e.m.), calculated from three independent simulations. (a) Distributions of end-to-end distance  $R$ , for chains comprising 30 – 360 bp, generated by Wang-Landau sampling. Melts are disallowed here, i.e.,  $\beta\overline{\Delta\mu} = \infty$ . (b) Distributions of end-to-end distance  $R$ , for 90 bp meltable wormlike chains for several values of the (renormalized) free energy of melting,  $\overline{\Delta\mu}$ .

Surprisingly, at the short chain lengths (less than  $\sim 90$  bp), a meltable chain becomes easier to cyclize with shorter chain length, reflected in a higher  $J$ -factor. This trend is rationalized through consideration of the enthalpic and entropic costs of cyclization. Cyclizing a short fully-hybridized DNA is so energetically costly (seen in the sharp drop in  $J$ -factor for the un-meltable chain (black curve)) that a short meltable chain essentially requires a melt near the middle to cyclize. Such a cyclized structure with a single melt will have an enthalpy largely independent of chain length, due to the high compliance of the melted region. However the entropic cost of cyclization decreases with decreasing chain length, so for short meltable chains the probability of cyclization increases with decreasing chain length.

This striking increase in  $J$ -factor for very short chains also occurs for the KWLC.<sup>21</sup> It does not appear in the spontaneous bend model with sufficiently rigid molten sections, nor in a model that fixes a particular bend an-

gle<sup>17</sup> (equivalent to the spontaneous bend model with infinite molten persistence length  $\ell_P^M$ ), because in these models cyclization of shorter chains still requires tighter bending and thus greater enthalpic cost. It also would not occur for the linear subelastic chain<sup>18</sup> as that model does not concentrate bending at one location. No experiments have yet probed the short lengths where this phenomenon is seen for the MWLC or KWLC.

We now compare with experiment these computational results for meltable wormlike chains with the stiffer renormalized melting free energies  $\overline{\Delta\mu} = 8 - 11 k_B T$  (Fig. 4b inset). Cloutier and Widom’s experimental cyclization rates<sup>16</sup> show significant scatter, presumably reflecting sequence heterogeneity in melt- and bend- propensities even among ‘random’ sequences (which exclude nucleosome positioning sequences known to have unusual bending characteristics). The meltable wormlike chain model predicts cyclization rates significantly lower than those inferred from experiment over the range of 90-120 bp,

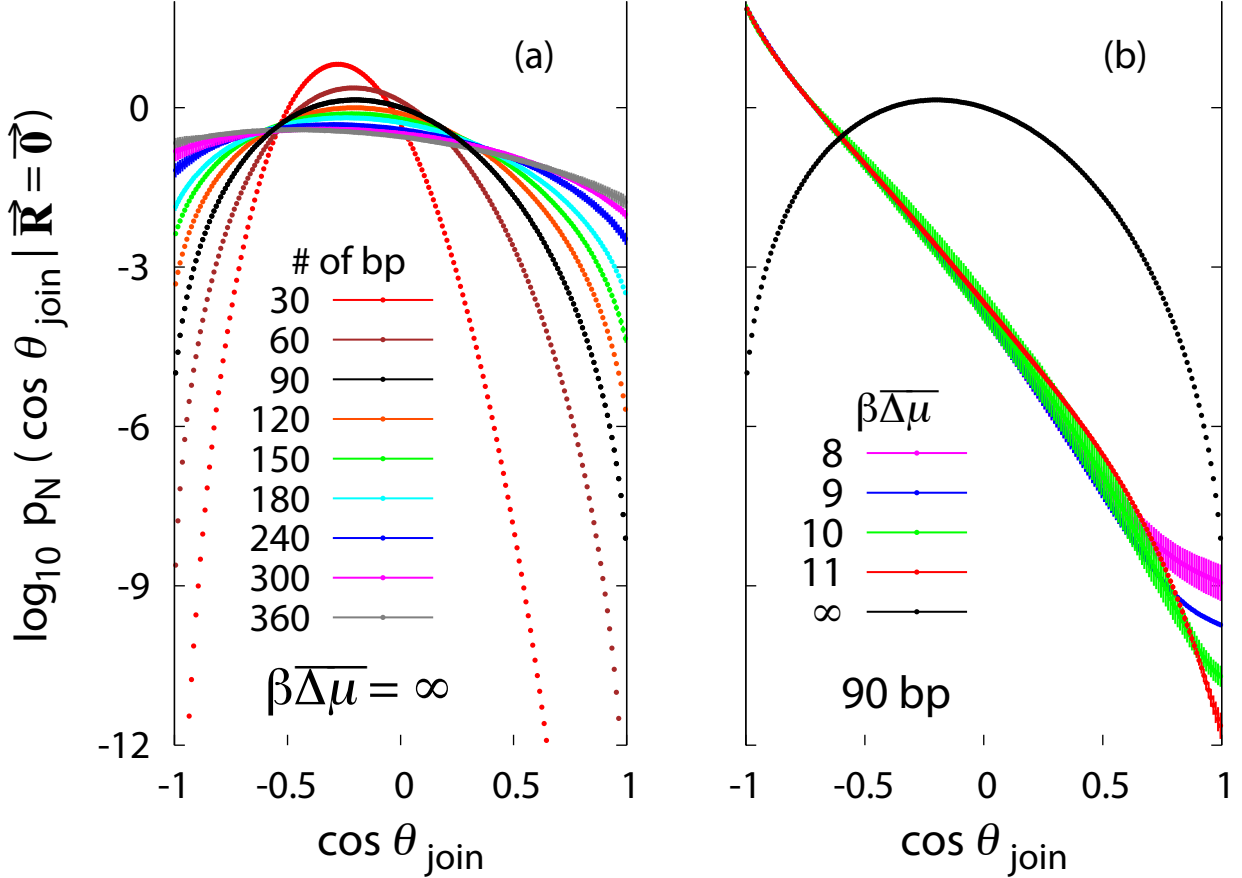


FIG. 3. Statistics of chain's initial orientation ( $\theta_0$ ), relative to that of its final link ( $\theta_N$ ), for cyclized melttable wormlike chains of varying contour length and melting thermodynamics. Error bars show  $\pm 2$  s.e.m. (a) Distributions of  $\cos \theta_{\text{join}}$ , for wormlike chains comprising 30 – 360 bp, generated by Wang-Landau sampling. Melts are disallowed here, i.e.,  $\beta \overline{\Delta \mu} = \infty$ . (b) Distributions of  $\cos \theta_{\text{join}}$  for 90 bp melttable wormlike chains for several values of the (renormalized) free energy of melting,  $\overline{\Delta \mu}$ .

which includes the majority of Cloutier and Widom's sequences. Melting excitations, when made thermodynamically consistent with bulk melting behavior, appear insufficient to explain their results.

The experimental results of Du, *et al.*<sup>17</sup> highlight the  $J$ -factor's strong helical dependence at these short lengths. Local maxima in  $J$  occur for chain lengths that are integer multiples of the helical repeat, thus requiring minimal over- or under-twisting to bring apposing ends into helical register. These chain lengths should offer the most straightforward comparison with the melttable wormlike chain model (which implicitly assumes perfect helical register). Throughout the range of  $N$  studied in these experiments, these maxima are slightly higher than corresponding  $J$ -factors of the melttable wormlike chain, or compare favorably with results for the smallest values of  $\Delta \mu$ . Given debate over the precise value for DNA persistence length within the range of  $\sim 45$ -53 nm, the experiments of Du, *et al.* appear to be consistent with a melttable wormlike chain model. But in this range of

uncertainty, their results are almost equally consistent with an unmelttable wormlike chain, which has similar  $J$ -factors to the melttable wormlike chain for molecules longer than  $\sim 120$  bp. Shorter chains would offer a much more discriminating test: for  $N \lesssim 50$  bp, the predicted  $J$ -factors of melttable and un-melttable chains differ enormously. Experiments on such molecules should unambiguously address the importance of thermally-excited melts in the severe bending of dsDNA.

## IX. TEMPERATURE DEPENDENCE

Enhanced cyclization of long DNA strands (thousands of bp) due to elevated temperature was established decades ago.<sup>37</sup> Yet there are few experiments (to our knowledge) probing temperature-variation of cyclization kinetics at the short lengths of interest here. Given the expected strong Arrhenius temperature-dependence of melting, such measurements should clarify the role

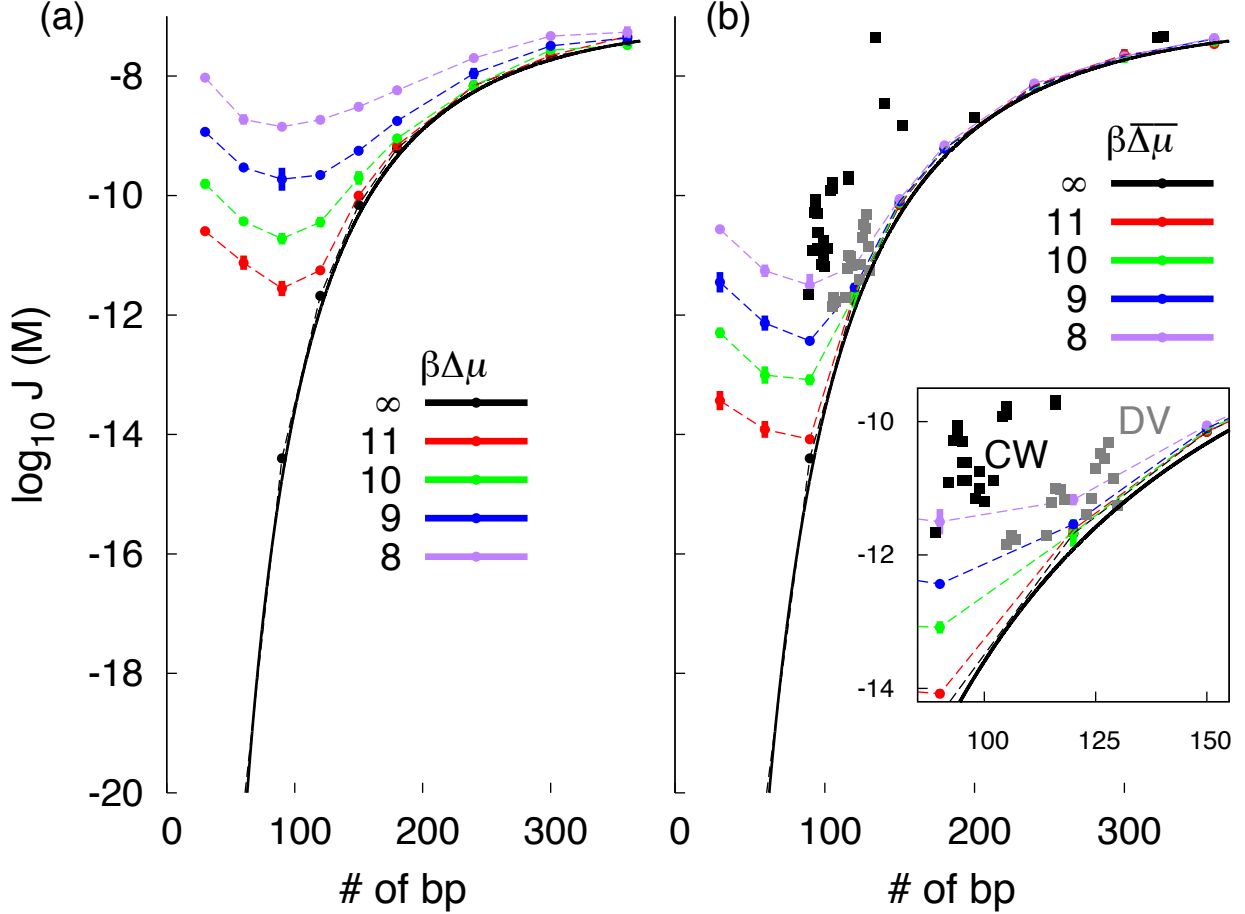


FIG. 4. Cyclization kinetics calculated for melttable wormlike chains of varying contour lengths and melting thermodynamics. Error bars show  $\pm 2$  s.e.m. (a)  $J$ -factor plotted as a function of contour length, for *bare* melting free energies  $\Delta\mu = 8 k_B T$  (magenta),  $9 k_B T$  (blue),  $10 k_B T$  (green),  $11 k_B T$  (red), and in the absence of melts,  $\Delta\mu = \infty$  (black). (b)  $J$ -factor plotted as a function of contour length, with *renormalized* melting free energies  $\overline{\Delta\mu} = 8 k_B T$  (magenta),  $9 k_B T$  (blue),  $10 k_B T$  (green),  $11 k_B T$  (red), and in the absence of melts,  $\overline{\Delta\mu} = \infty$  (black). Solid black line shows predictions from the approximate analytic theory of Shimada and Yamakawa (in the absence of melts).<sup>36</sup> Black squares show experimental results of Cloutier and Widom;<sup>16</sup> gray squares those of Du, *et al.*<sup>17</sup>

of dehybridization or even explain discrepancies between different experiments. Here we scrutinize these possibilities through the thermal sensitivity of an MWLC.

In this section we present numerical results for  $J$ -factors as a function of  $T$ . We assume (i) temperature-independent bending rigidities  $k_B T \ell_P^H$  and  $k_B T \ell_P^M$ ; and (ii) a constant value of the apparent bending rigidity  $k_B T \ell_P$ , a weighted combination of hybridized and melted bending rigidities that is dominated by the former. Transient electric birefringence experiments in the range of 20–43°C<sup>38</sup> support this latter assumption.

Fig. 5 shows computed  $J$ -factors, for the case of  $\overline{\Delta\mu} = 10 k_B T$ , at several different temperatures. We find that a change of 10°C has very little impact on  $J$  for all chain lengths examined. In fact varying temperature by as much as 30°C never altered  $J$  by more than a factor of three, for all lengths examined. Experiments spanning 23–42°C found similar temperature-sensitivity

for a 200 bp segment of  $\lambda$  DNA and a 116 bp chain with sequence chosen to minimize melting, but somewhat greater temperature-sensitivity in a 116 bp chain with sequence designed to accentuate melting.<sup>25</sup> Since the temperatures in the respective experiments of Du *et al.* and Cloutier and Widom differ by only nine degrees C, temperature variation is unlikely to explain differences in their findings.

## X. CONCLUSION

In this work we have explored the mutual influence of distinct degrees of freedom characterizing the microscopic state of a chain molecule. Our results highlight how the presence or absence of one fluctuating variable shapes the effective forces on another. From a practical perspective, these results caution that grafting an

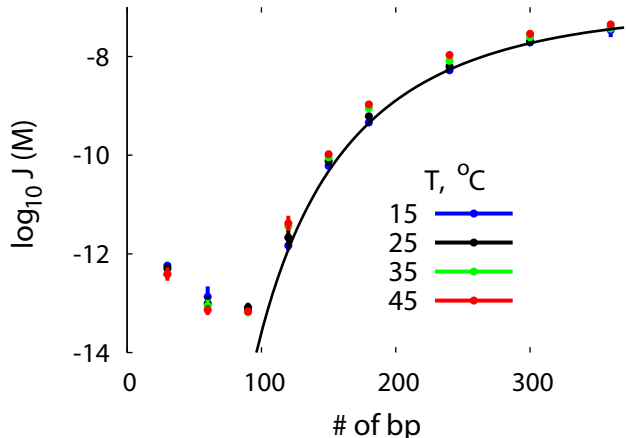


FIG. 5. Temperature dependence of cyclization kinetics for meltable wormlike chains.  $J$ -factors are plotted as functions of contour length for  $T = 15, 25, 35$ , and  $45^\circ\text{C}$ . A moderate melt enthalpy,  $\Delta\mu = 10 k_B T$ , was used in all cases. Error bars show  $\pm 2$  s.e.m. Solid black line shows predictions from the approximate analytic theory of Shimada and Yamakawa (in the absence of melts).<sup>36</sup>

additional degree of freedom onto a model changes the apparent parameters governing the fluctuations of other degrees of freedom.

These concepts find a concrete example in models of DNA flexibility that incorporate regions of thermally-enhanced flexibility. It seems immediately clear that introducing the possibility of flexible excitations to a polymer will reduce its apparent persistence length, hence the invocation of melts in attempts to understand anomalously high cyclization rates. Somewhat more subtly, adding bending degrees of freedom (that differ in their natural fluctuations between hybridized and melted sections of the chain) to each site in a lattice model for melting can substantially renormalize the free energy of dehybridization. For the particular parameters used in our simulations ( $\ell_P^H = 50$  nm,  $\ell_P^M = 2.5$  nm,  $d = 1$  nm), the bending-induced increase in apparent melt free energy amounts to  $3.7 k_B T$ .

Based on this distinction between bare and apparent melt free energies, we develop a thermodynamically consistent treatment of the meltable wormlike chain model that properly accounts for the greater entropy of more flexible melted sections, and thus permits accurate parameterization from melting experiments. For sufficiently short DNA chains, the conformations that lead to cyclization are exceedingly rare and thus their equilibrium probability is greatly facilitated by the transiently enhanced flexibility of thermally-excited melts, even at the energetic cost of disrupting DNA basepairing and stacking interactions. However, the proper parameterization of melting energies in the range  $\Delta\mu = 8 - 11 k_B T$  pushes the onset of such melt-enhanced cyclization to significantly shorter lengths, so much so that the cy-

clization propensity of a meltable chain is practically indistinguishable from an un-meltable chain for chain lengths greater than  $\sim 120$  bp. Predictions for meltable and un-meltable chains diverge significantly around  $\sim 80$  bp, suggesting that experiments on even shorter chains would distinguish between models. Indeed at the shortest lengths, meltable chains even show an *increase* in cyclization propensity with decreasing chain length, due to entropic effects.

We have also demonstrated that coarse-grained modeling with Wang-Landau sampling can be a useful methodology for calculating  $J$ -factors for short DNA chains. This basic approach could find further application in exploring additional aspects of DNA bending elasticity, of great relevance for DNA compaction and protein-DNA interactions. Elaborations of the MWLC model may be required for these purposes. Our existing computational framework could easily incorporate the dependence of bending elasticity on salt concentrations and DNA sequence, through altered free energies of melting<sup>31</sup> and altered flexibility of a given melted or hybridized section.<sup>39,40</sup> Other possible straightforward amendments include more detailed melt thermodynamics (for example incorporating stacking interactions,<sup>41</sup> a specific heat term<sup>42</sup> or spatial and dynamical information<sup>43</sup>), volume exclusion, and explicit representation of DNA helicity.<sup>30</sup> Nevertheless, the meltable WLC still provides a parsimonious model lending important insights regarding the effect on DNA cyclization behavior of greater flexibility conferred by thermally-excited defects.

## ACKNOWLEDGMENTS

Steve Whitelam provided many useful discussions. DS acknowledges support from a National Science Foundation Graduate Research fellowship.

- <sup>1</sup>C. Bustamante, S. B. Smith, J. T. Liphardt, and D. Smith, "Single-molecule studies of DNA mechanics," *Curr. Opin. Struct. Biol.*, **10**, 279 (2000).
- <sup>2</sup>S. B. Smith, Y. Cui, and C. Bustamante, "Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules," *Science*, **271**, 795 (1996).
- <sup>3</sup>C. Bustamante, Z. Bryant, and S. B. Smith, "Ten years of tension: single-molecule DNA mechanics," *Nature*, **421**, 423 (2003).
- <sup>4</sup>D. Crothers, J. Drak, J. Kahn, and S. Levene, "DNA bending, flexibility, and helical repeat by cyclization kinetics," *Meth. Enzym.*, **212**, 3 (1992).
- <sup>5</sup>J. Perez-Martin and V. de Lorenzo, "Clues and consequences of DNA bending in transcription," *Ann. Rev. Microbiol.*, **51**, 593 (1997).
- <sup>6</sup>T. J. Richmond and C. A. Davey, "The structure of DNA in the nucleosome core," *Nature*, **423**, 145 (2003).
- <sup>7</sup>P. A. Rice, S. W. Yang, K. Mizuuchi, and H. A. Nash, "Crystal structure of an IHF-DNA complex: A protein-induced DNA U-turn," *Cell*, **87**, 1295 (1996).
- <sup>8</sup>H. G. Garcia, P. Grayson, L. Han, M. Inamdar, J. Kondev, P. C. Nelson, R. Phillips, J. Widom, and P. A. Wiggins, "Biological consequences of tightly bent DNA: The other life of a macromolecular celebrity," *Biopolymers*, **85**, 115 (2007).

- <sup>9</sup>C. A. Mirkin, R. Letsinger, R. Mucic, and J. Storhoff, "A DNA-based method for rationally assembling nanoparticles into macroscopic materials," *Nature*, **382**, 607 (1996).
- <sup>10</sup>A. Fu, C. M. Micheel, J. Cha, H. Chang, H. Yang, and A. P. Alivisatos, "Discrete nanostructures of quantum dots/Au with DNA," *J. Am. Chem. Soc.*, **126**, 10832 (2004).
- <sup>11</sup>S. Claridge, A. J. Mastroianni, Y. Au, H. Liang, C. M. Micheel, J. M. J. Frechet, and A. P. Alivisatos, "Enzymatic ligation creates discrete multinanoparticle building blocks for self-assembly," *J. Am. Chem. Soc.*, **130**, 9598 (2008).
- <sup>12</sup>H. Dietz, S. M. Douglas, and W. M. Shih, "Folding DNA into twisted and curved nanoscale shapes," *Science*, **325**, 725 (2009).
- <sup>13</sup>S. Y. Park, A. K. R. Lytton-Jean, B. Lee, S. Weigand, G. C. Schatz, and C. A. Mirkin, "DNA-programmable nanoparticle crystallization," *Nature*, **451**, 553 (2008).
- <sup>14</sup>P. A. Wiggins, T. van der Heijden, F. Moreno-Herrero, A. J. Spakowitz, R. S. Phillips, J. Widom, C. Dekker, and P. C. Nelson, "High flexibility of DNA on short length scales probed by atomic force microscopy," *Nat. Nanotech.*, **1**, 137 (2006).
- <sup>15</sup>C. Yuan, H. Chen, X. W. Lou, and L. A. Archer, "DNA bending stiffness on small length scales," *Phys. Rev. Lett.*, **100**, 018102 (2008).
- <sup>16</sup>T. E. Cloutier and J. Widom, "Spontaneous sharp bending of double-stranded DNA," *Mol. Cell.*, **14**, 355 (2004).
- <sup>17</sup>Q. Du, C. Smith, N. Shiffeldrim, M. Vologodskaya, and A. V. Vologodskii, "Cyclization of short DNA fragments and bending fluctuations of the double helix," *Proc. Natl. Acad. Sci. USA*, **102**, 5397 (2005).
- <sup>18</sup>P. A. Wiggins and P. C. Nelson, "Generalized theory of semiflexible polymers," *Phys. Rev. E*, **73**, 031906 (2006).
- <sup>19</sup>J. Yan and J. F. Marko, "Localized single-stranded bubble mechanism for cyclization of short double helix DNA," *Phys. Rev. Lett.*, **93**, 108108 (2004).
- <sup>20</sup>J. Yan, R. Kawamura, and J. F. Marko, "Statistics of loop formation along double helix DNAs," *Phys. Rev. E*, **71**, 061905 (2005).
- <sup>21</sup>P. A. Wiggins, R. S. Phillips, and P. C. Nelson, "Exact theory of kinkable elastic polymers," *Phys. Rev. E*, **71**, 021909 (2005).
- <sup>22</sup>M. D. Wang, H. Yin, R. Landick, J. Gelles, and S. M. Block, "Stretching DNA with optical tweezers," *Biophys. J.*, **72**, 1335 (1997).
- <sup>23</sup>D. Shore, J. Langowski, and R. L. Baldwin, "DNA flexibility studied by covalent closure of short fragments into circles," *Proc. Natl. Acad. Sci. USA*, **78**, 4833 (1981).
- <sup>24</sup>W. H. Taylor and P. J. Hagerman, "Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure. II. NaCl-dependence of DNA flexibility and helical repeat," *J. Mol. Biol.*, **212**, 363 (1990).
- <sup>25</sup>R. A. Forties, R. Bundschuh, and M. G. Poirier, "The flexibility of locally melted DNA," *Nuc. Acids Res.*, **37**, 4580 (2009).
- <sup>26</sup>J. P. Peters and L. J. M. III, "DNA curvature and flexibility in vitro and in vivo," *Q. Rev. Biophys.*
- <sup>27</sup>R. R. Cheng, T. Uzawa, K. W. Plaxco, and D. E. Makarov, "Universality in the timescales of internal loop formation in unfolded proteins and single-stranded oligonucleotides," *Biophys. J.*, **99**, 3959 (2010).
- <sup>28</sup>B. H. Zimm and J. K. Bragg, "Theory of the phase transition between helix and random coil in polypeptide chains," *J. Chem. Phys.*, **31**, 526 (1959).
- <sup>29</sup>J. Palmeri, M. Manghi, and N. Destainville, "Thermal denaturation of fluctuating DNA driven by bending entropy," *Phys. Rev. Lett.*, **99**, 088103 (2007).
- <sup>30</sup>A. J. Mastroianni, D. A. Sivak, P. L. Geissler, and A. P. Alivisatos, "Probing the conformational distributions of subpersistence length DNA," *Biophys. J.*, **97**, 1408 (2009).
- <sup>31</sup>J. SantaLucia, "A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics," *Proc. Nat. Acad. Sci. USA*, **95**, 1460 (1998).
- <sup>32</sup>D. P. Landau, S. Tsai, and M. Exler, "A new approach to monte carlo simulations in statistical physics: Wang-landau sampling," *Am J Phys*, **72**, 1294 (2004).
- <sup>33</sup>J. F. Allemand, S. Cocco, N. Douarche, and G. Lia, "Loops in DNA: an overview of experimental and theoretical approaches," *Eur. Phys. J. E*, **19**, 293 (2006).
- <sup>34</sup>H. Yamakawa and W. Stockmayer, "Statistical mechanics of wormlike chains. 2. excluded volume effects," *J. Chem. Phys.*, **57**, 2843 (1972).
- <sup>35</sup>M. C. Murphy, I. Rasnik, W. Cheng, T. M. Lohman, and T. Ha, "Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy," *Biophys. J.*, **86**, 2530 (2004).
- <sup>36</sup>J. Shimada and H. Yamakawa, "Ring-closure probabilities for twisted wormlike chains. application to DNA," *Macromol.*, **17**, 689 (1984).
- <sup>37</sup>J. E. Mertz and R. W. Davis, "Cleavage of DNA by r1 restriction endonuclease generates cohesive ends," *Proc. Natl. Acad. Sci. USA*, **69**, 3370 (1972).
- <sup>38</sup>Y. Lu, B. Weers, and N. C. Stellwagen, "DNA persistence length revisited," *Biopol.*, **61**, 261 (2001).
- <sup>39</sup>C. Baumann, S. B. Smith, V. A. Bloomfield, and C. Bustamante, "Ionic effects on the elasticity of single DNA molecules," *Proc. Natl. Acad. Sci. USA*, **94**, 6185 (1997).
- <sup>40</sup>W. K. Olson, A. A. Gorin, X.-J. Lu, L. Hock, and V. B. Zhurkin, "DNA sequence-dependent deformability deduced from protein-DNA crystal complexes," *Proc. Natl. Acad. Sci. USA*, **95**, 11163 (1998).
- <sup>41</sup>A. Krueger, E. Protozanova, and M. D. Frank-Kamenetskii, "Sequence-dependent base pair opening in DNA double helix," *Biophys. J.*, **90**, 3091 (2006).
- <sup>42</sup>M. C. Williams, J. R. Wenner, I. Rouzina, and V. A. Bloomfield, "Entropy and heat capacity of DNA melting from temperature dependence of single molecule stretching," *Biophys. J.*, **80**, 1932 (2001).
- <sup>43</sup>T. Dauxois, M. Peyrard, and A. R. Bishop, "Dynamics and thermodynamics of a nonlinear model for DNA denaturation," *Phys. Rev. E*, **47**, 684 (1993).